PCT

RAW SEQUENCE LISTING DATE: 07/16/2004
PATENT APPLICATION: **US/10/501,071** TIME: 16:37:40

Input Set : **A:\Sequence Listing.ST25.txt**
Output Set: **N:\CRF4\07162004\J501071.raw**

```
     5 <110> APPLICANT: University of Newcastle Upon Tyne
     9 <120> TITLE OF INVENTION: Fusion Proteins
    13 <130> FILE REFERENCE: 43952/JMD/MAR
C--> 17 <140> CURRENT APPLICATION NUMBER: US/10/501,071
C--> 17 <141> CURRENT FILING DATE: 2004-07-09
    17 <150> PRIOR APPLICATION NUMBER: GB 0200689.8
    19 <151> PRIOR FILING DATE: 2002-01-10
    23 <160> NUMBER OF SEQ ID NOS: 61
    27 <170> SOFTWARE: PatentIn version 3.1
    31 <210> SEQ ID NO: 1
    33 <211> LENGTH: 9
    35 <212> TYPE: PRT
    37 <213> ORGANISM: Artificial Sequence
    41 <220> FEATURE:
    43 <223> OTHER INFORMATION: Ala3-His6 tail
    45 <400> SEQUENCE: 1
    47 Ala Ala Ala His His His His His His
    48 1               5
    51 <210> SEQ ID NO: 2
    53 <211> LENGTH: 25
    55 <212> TYPE: PRT
    57 <213> ORGANISM: Escherichia coli
    61 <400> SEQUENCE: 2
    63 Met Asn Met Lys Lys Leu Ala Thr Leu Val Ser Ala Val Ala Leu Ser
    64 1               5                  10                  15
    67 Ala Thr Val Ser Ala Asn Ala Met Ala
    68             20                  25
    71 <210> SEQ ID NO: 3
    73 <211> LENGTH: 5
    75 <212> TYPE: PRT
    77 <213> ORGANISM: Artificial Sequence
    81 <220> FEATURE:
    83 <223> OTHER INFORMATION: Cleavage site for enterokinase
    85 <400> SEQUENCE: 3
    87 Asp Asp Asp Asp Lys
    88 1               5
    91 <210> SEQ ID NO: 4
    93 <211> LENGTH: 4
    95 <212> TYPE: PRT
    97 <213> ORGANISM: Artificial Sequence
   101 <220> FEATURE:
   103 <223> OTHER INFORMATION: Cleavage site for thrombin
   105 <400> SEQUENCE: 4
```

ENTERED

RAW SEQUENCE LISTING DATE: 07/16/2004
PATENT APPLICATION: US/10/501,071 TIME: 16:37:40

Input Set : A:\Sequence Listing.ST25.txt
Output Set: N:\CRF4\07162004\J501071.raw

107 Leu Val Pro Arg
108 1
111 <210> SEQ ID NO: 5
113 <211> LENGTH: 4
115 <212> TYPE: PRT
117 <213> ORGANISM: Artificial Sequence
121 <220> FEATURE:
123 <223> OTHER INFORMATION: Cleavage site for factor Xa
125 <400> SEQUENCE: 5
127 Ile Glu Gly Arg
128 1
131 <210> SEQ ID NO: 6
133 <211> LENGTH: 4
135 <212> TYPE: PRT
137 <213> ORGANISM: Artificial Sequence
141 <220> FEATURE:
143 <223> OTHER INFORMATION: 4xHis tag
145 <400> SEQUENCE: 6
147 His His His His
148 1
151 <210> SEQ ID NO: 7
153 <211> LENGTH: 5
155 <212> TYPE: PRT
157 <213> ORGANISM: Artificial Sequence
161 <220> FEATURE:
163 <223> OTHER INFORMATION: 5xHis tag
165 <400> SEQUENCE: 7
167 His His His His His
168 1 5
171 <210> SEQ ID NO: 8
173 <211> LENGTH: 6
175 <212> TYPE: PRT
177 <213> ORGANISM: Artificial Sequence
181 <220> FEATURE:
183 <223> OTHER INFORMATION: 6xHis tag
185 <400> SEQUENCE: 8
187 His His His His His His
188 1 5
191 <210> SEQ ID NO: 9
193 <211> LENGTH: 7
195 <212> TYPE: PRT
197 <213> ORGANISM: Artificial Sequence
201 <220> FEATURE:
203 <223> OTHER INFORMATION: 7xHis tag
205 <400> SEQUENCE: 9
207 His His His His His His His
208 1 5
211 <210> SEQ ID NO: 10
213 <211> LENGTH: 8

RAW SEQUENCE LISTING DATE: 07/16/2004
PATENT APPLICATION: US/10/501,071 TIME: 16:37:40

Input Set : A:\Sequence Listing.ST25.txt
Output Set: N:\CRF4\07162004\J501071.raw

215 <212> TYPE: PRT
217 <213> ORGANISM: Artificial Sequence
221 <220> FEATURE:
223 <223> OTHER INFORMATION: 8xHis tag
225 <400> SEQUENCE: 10
227 His His His His His His His His
228 1 5
231 <210> SEQ ID NO: 11
233 <211> LENGTH: 9
235 <212> TYPE: PRT
237 <213> ORGANISM: Artificial Sequence
241 <220> FEATURE:
243 <223> OTHER INFORMATION: 9xHis tag
245 <400> SEQUENCE: 11
247 His His His His His His His His His
248 1 5
251 <210> SEQ ID NO: 12
253 <211> LENGTH: 10
255 <212> TYPE: PRT
257 <213> ORGANISM: Artificial Sequence
261 <220> FEATURE:
263 <223> OTHER INFORMATION: 10xHis tag
265 <400> SEQUENCE: 12
267 His His His His His His His His His His
268 1 5 10
271 <210> SEQ ID NO: 13
273 <211> LENGTH: 93
275 <212> TYPE: PRT
277 <213> ORGANISM: Escherichia coli
281 <400> SEQUENCE: 13
283 Asn Asn Gly Ala Ser Gly Ala Asp Ile Asn Asn Tyr Ala Gly Gln Ile
284 1 5 10 15
287 Lys Ser Ala Ile Glu Ser Lys Phe Tyr Asp Ala Ser Ser Tyr Ala Gly
288 20 25 30
291 Lys Thr Cys Thr Leu Arg Ile Lys Leu Ala Pro Asp Gly Met Leu Leu
292 35 40 45
295 Asp Ile Lys Pro Glu Gly Gly Asp Pro Ala Leu Cys Gln Ala Ala Leu
296 50 55 60
299 Ala Ala Ala Lys Leu Ala Lys Ile Pro Lys Pro Pro Ser Gln Ala Val
300 65 70 75 80
303 Tyr Glu Val Phe Lys Asn Ala Pro Leu Asp Phe Lys Pro
304 85 90
307 <210> SEQ ID NO: 14
309 <211> LENGTH: 348
311 <212> TYPE: PRT
313 <213> ORGANISM: Artificial Sequence
317 <220> FEATURE:
319 <223> OTHER INFORMATION: TolA-BCL fusion protein
321 <400> SEQUENCE: 14

RAW SEQUENCE LISTING DATE: 07/16/2004
PATENT APPLICATION: US/10/501,071 TIME: 16:37:40

Input Set : A:\Sequence Listing.ST25.txt
Output Set: N:\CRF4\07162004\J501071.raw

```
323 Met His His His His His His Ser Ser Asn Asn Gly Ala Ser Gly Ala
324 1               5                   10                  15
327 Asp Ile Asn Asn Tyr Ala Gly Gln Ile Lys Ser Ala Ile Glu Ser Lys
328             20                  25                  30
331 Phe Tyr Asp Ala Ser Ser Tyr Ala Gly Lys Thr Cys Thr Leu Arg Ile
332         35                  40                  45
335 Lys Leu Ala Pro Asp Gly Met Leu Leu Asp Ile Lys Pro Glu Gly Gly
336     50                  55                  60
339 Asp Pro Ala Leu Cys Gln Ala Ala Leu Ala Ala Ala Lys Leu Ala Lys
340 65                  70                  75                  80
343 Ile Pro Lys Pro Pro Ser Gln Ala Val Tyr Glu Val Phe Lys Asn Ala
344                 85                  90                  95
347 Pro Leu Asp Phe Lys Pro Gly Gly Gly Ser Gly Ser Leu Val Pro Arg
348             100                 105                 110
351 Gly Ser Arg Pro Ser Gln Ser Asn Arg Glu Leu Val Val Asp Phe Leu
352         115                 120                 125
355 Ser Tyr Lys Leu Ser Gln Lys Gly Tyr Ser Trp Ser Gln Phe Ser Asp
356     130                 135                 140
359 Val Glu Glu Asn Arg Thr Glu Ala Pro Glu Gly Thr Glu Ser Glu Met
360 145                 150                 155                 160
363 Glu Thr Pro Ser Ala Ile Asn Gly Asn Pro Ser Trp His Leu Ala Asp
364                 165                 170                 175
367 Ser Pro Ala Val Asn Gly Ala Thr Ala His Ser Ser Ser Leu Asp Ala
368             180                 185                 190
371 Arg Glu Val Ile Pro Met Ala Ala Val Lys Gln Ala Leu Arg Glu Ala
372         195                 200                 205
375 Gly Asp Glu Phe Glu Leu Arg Tyr Arg Arg Ala Phe Ser Asp Leu Thr
376     210                 215                 220
379 Ser Gln Leu His Ile Thr Pro Gly Thr Ala Tyr Gln Ser Phe Glu Gln
380 225                 230                 235                 240
383 Val Val Asn Glu Leu Phe Arg Asp Gly Val Asn Trp Gly Arg Ile Val
384                 245                 250                 255
387 Ala Phe Phe Ser Phe Gly Gly Ala Leu Cys Val Glu Ser Val Asp Lys
388             260                 265                 270
391 Glu Met Gln Val Leu Val Ser Arg Ile Ala Ala Trp Met Ala Thr Tyr
392         275                 280                 285
395 Leu Asn Asp His Leu Glu Pro Trp Ile Gln Glu Asn Gly Gly Trp Asp
396     290                 295                 300
399 Thr Phe Val Glu Leu Tyr Gly Asn Asn Ala Ala Ala Glu Ser Arg Lys
400 305                 310                 315                 320
403 Gly Gln Glu Arg Phe Asn Arg Trp Phe Leu Thr Gly Met Thr Val Ala
404                 325                 330                 335
407 Gly Val Val Leu Leu Gly Ser Leu Phe Ser Arg Lys
408             340                 345
411 <210> SEQ ID NO: 15
413 <211> LENGTH: 236
415 <212> TYPE: PRT
417 <213> ORGANISM: Artificial Sequence
421 <220> FEATURE:
```

RAW SEQUENCE LISTING DATE: 07/16/2004
PATENT APPLICATION: US/10/501,071 TIME: 16:37:40

Input Set : A:\Sequence Listing.ST25.txt
Output Set: N:\CRF4\07162004\J501071.raw

423 <223> OTHER INFORMATION: TolA-BCL fusion protein after thrombin cleavage
425 <400> SEQUENCE: 15
427 Gly Ser Arg Pro Ser Gln Ser Asn Arg Glu Leu Val Val Asp Phe Leu
428 1 5 10 15
431 Ser Tyr Lys Leu Ser Gln Lys Gly Tyr Ser Trp Ser Gln Phe Ser Asp
432 20 25 30
435 Val Glu Glu Asn Arg Thr Glu Ala Pro Glu Gly Thr Glu Ser Glu Met
436 35 40 45
439 Glu Thr Pro Ser Ala Ile Asn Gly Asn Pro Ser Trp His Leu Ala Asp
440 50 55 60
443 Ser Pro Ala Val Asn Gly Ala Thr Ala His Ser Ser Ser Leu Asp Ala
444 65 70 75 80
447 Arg Glu Val Ile Pro Met Ala Ala Val Lys Gln Ala Leu Arg Glu Ala
448 85 90 95
451 Gly Asp Glu Phe Glu Leu Arg Tyr Arg Arg Ala Phe Ser Asp Leu Thr
452 100 105 110
455 Ser Gln Leu His Ile Thr Pro Gly Thr Ala Tyr Gln Ser Phe Glu Gln
456 115 120 125
459 Val Val Asn Glu Leu Phe Arg Asp Gly Val Asn Trp Gly Arg Ile Val
460 130 135 140
463 Ala Phe Phe Ser Phe Gly Gly Ala Leu Cys Val Glu Ser Val Asp Lys
464 145 150 155 160
467 Glu Met Gln Val Leu Val Ser Arg Ile Ala Ala Trp Met Ala Thr Tyr
468 165 170 175
471 Leu Asn Asp His Leu Glu Pro Trp Ile Gln Glu Asn Gly Gly Trp Asp
472 180 185 190
475 Thr Phe Val Glu Leu Tyr Gly Asn Asn Ala Ala Ala Glu Ser Arg Lys
476 195 200 205
479 Gly Gln Glu Arg Phe Asn Arg Trp Phe Leu Thr Gly Met Thr Val Ala
480 210 215 220
483 Gly Val Val Leu Leu Gly Ser Leu Phe Ser Arg Lys
484 225 230 235
487 <210> SEQ ID NO: 16
489 <211> LENGTH: 115
491 <212> TYPE: PRT
493 <213> ORGANISM: Artificial Sequence
497 <220> FEATURE:
499 <223> OTHER INFORMATION: Tagged TolAIII region of pTol vectors
501 <220> FEATURE:
503 <221> NAME/KEY: MISC_FEATURE
505 <222> LOCATION: (107)..(111)
507 <223> OTHER INFORMATION: Xaa residues represent cleavage sites DDDDK (SEQ ID NO: 3), LVPR
508 (SEQ ID NO: 4; no Xaa at position 111) or IEGR (SEQ ID NO: 5; no
509 Xaa at position 111)
513 <400> SEQUENCE: 16
515 Met His His His His His His Ser Ser Asn Asn Gly Ala Ser Gly Ala
516 1 5 10 15
519 Asp Ile Asn Asn Tyr Ala Gly Gln Ile Lys Ser Ala Ile Glu Ser Lys
520 20 25 30

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 07/16/2004
PATENT APPLICATION: US/10/501,071 TIME: 16:37:41

Input Set : A:\Sequence Listing.ST25.txt
Output Set: N:\CRF4\07162004\J501071.raw

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:16; Xaa Pos. 107,108,109,110,111
Seq#:22; Xaa Pos. 14,15
Seq#:23; Xaa Pos. 13,14
Seq#:24; Xaa Pos. 13,14

VERIFICATION SUMMARY DATE: 07/16/2004
PATENT APPLICATION: **US/10/501,071** TIME: 16:37:41

Input Set : **A:\Sequence Listing.ST25.txt**
Output Set: **N:\CRF4\07162004\J501071.raw**

L:17 M:270 C: Current Application Number differs, Replaced Current Application No
L:17 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:539 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16 after pos.:96
L:667 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22 after pos.:0
L:701 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:23 after pos.:0
L:731 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:24 after pos.:0

STATISTICS SUMMARY DATE: 07/16/2004
PATENT APPLICATION: **US/10/501,071** TIME: 16:37:41

Input Set : **A:\Sequence Listing.ST25.txt**
Output Set: **N:\CRF4\07162004\J501071.raw**

Application Serial Number: US/10/501,071
Alpha or Numeric or Xml: Numeric
Application Class:
Application File Date: 07-09-2004
Art Unit: PCT
Software Application: PatentIN3.1
Total Number of Sequences: 61
Total Nucleotides: 1032
Total Amino Acids: 2629
Number of Errors: 0
Number of Warnings: 4
Number of Corrections: 2

MESSAGE SUMMARY

270 C: 1 (Current Application Number differs)
271 C: 1 (Current Filing Date differs)
341 W: 4 ((46) "n" or "Xaa" used)